

1 Arcface

Softmax loss:

$$L_1 = -\frac{1}{N} \sum_{i=1}^N \log \frac{\exp(W_{y_i} \cdot x_i + b_{y_i})}{\sum_{j=1}^n \exp(W_j \cdot x_i + b_j)}$$

where

Table 1: Explanation	
N	batch size
n	number of classes
$x_i \in \mathbb{R}^d$	i -th sample's embedding
y_i	Its class
$W \in \mathbb{R}^{d \times n}$	Weight Matrix
$b_j \in \mathbb{R}^n$	Biased term

L_2 normalization on W, x_i